

# Báo Cáo Bài Tập Tuần BigData

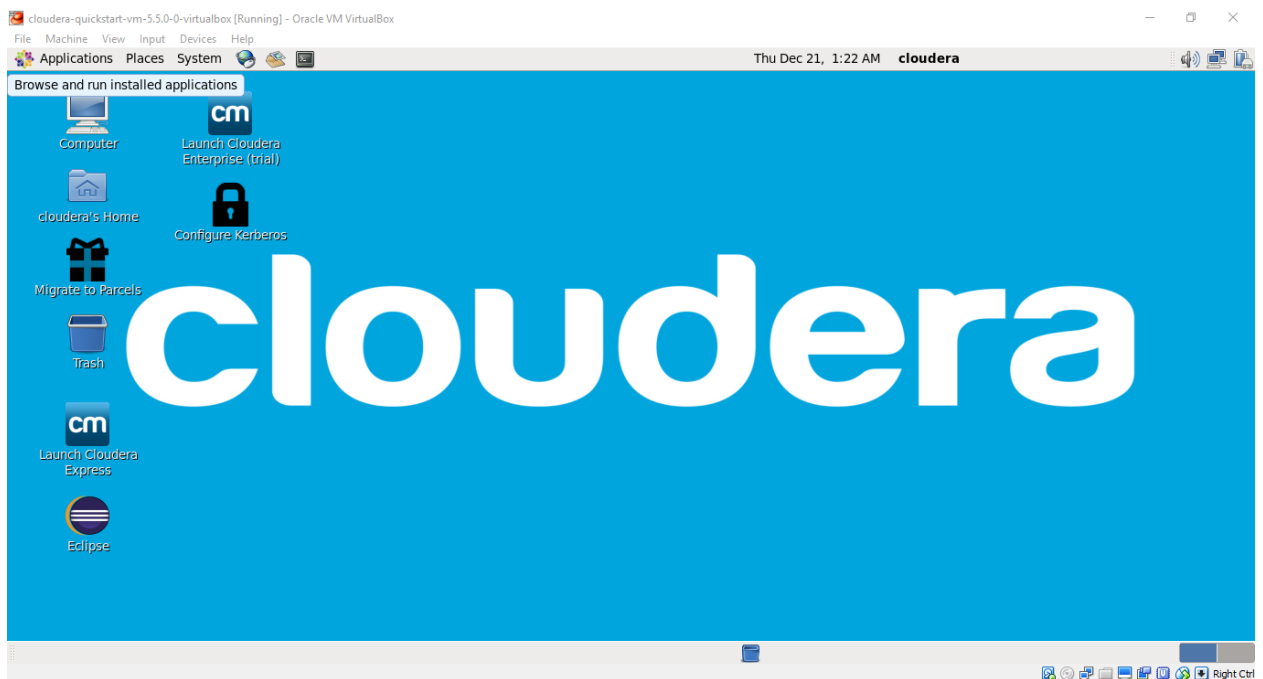
## Bài Tập Hadoop-MapReduce

Nhóm sinh viên:

Trần Văn Thành – MSSV 20133561

Phan Xuân Đức – MSSV 20146212

### 1. Minh chứng cài đặt Hadoop.



### 2. Mô tả thuật toán Mapreduce để giải bài tập "People You Might Know"

Bước **map** đầu tiên tạo ra cặp người dùng hoặc là bạn với nhau hoặc có bạn bè chung. Nếu người dùng đã là bạn bè, nó xuất ra một bộ  $((user, friend), 0)$ , và nếu người dùng có bạn bè chung, nó sẽ xuất ra một bộ  $(user1, user2), 1)$ .

Bước **reduce** đầu tiên đếm số bạn bè chung của mỗi cặp người dùng. Các phím là cặp của người dùng, và các giá trị là 0 và 1 được tạo ra trong bước bản đồ. Bộ giảm tốc finds (a) tổng các giá trị và (b) sản phẩm của các giá trị. Nếu sản phẩm của các giá trị là 0, sau đó người dùng đã được bạn bè, vì vậy chúng tôi không nên giới thiệu những người dùng với nhau.

Nếu không, chúng ta sẽ xuất ra bộ tuple ((user1, user2), numCommonFriends).

Bước **map** thứ hai ánh xạ bộ tuple ((user1, user2), numCommonFriends) vào tuple (user1, (user2, numCommonFriends)).

Bước **giảm** thứ hai lấy các giá trị đó và trích ra 10 giá trị của user2 có những người bạn thông dụng nhất với user1.

### 3. Source Code chương trình cài đặt thuật toán trong mục 2 trên nền tảng Hadoop.

File : FriendRecommendation.java.

[https://github.com/thanhtv2/Hadoop\\_MapReduce](https://github.com/thanhtv2/Hadoop_MapReduce)